

# 15

## 真核生物基因的表达调控

**真**核生物具有由膜围成的细胞核,其基因的转录和翻译分别在细胞核和细胞质中进行,将基因的表达过程非常明显地分成了两个阶段。此外,真核基因本身结构较为复杂,基因转录模板——**DNA** 又是与组蛋白等构成染色质的基本单位——核小体。因此,与原核生物相比,真核生物基因表达是处在一个非常纷繁复杂的控制系统中,表达通路的每一步都受到严格的调节。尽管多细胞真核生物各种类型的细胞都携带着相同遗传信息的基因组,但它们基因表达的程序和状况却不尽相同。各类细胞中特异基因的表达和相应的调控,也正是机体生长和发育分化的重要前提条件。有关基因表达调控的探索一直是基因分子生物学的一个前沿领域,所获得的研究成果使我们不仅能更好地认识这一复杂而有序的过程和基因发挥自身功能的意义,且有助于深刻地认识和阐释许多重要的生命现象。真核生物中编码蛋白质的基因不仅种类繁多,且结构复杂,它们的表达与高等真核生物的发育调控和形态建成息息相关。因此我们在真核生物基因表达调控的讨论中将侧重于 **RNA** 聚合酶 II 所转录的这类基因的表达调控。

## 15.1 真核生物基因转录水平调节

### 15.1.1 真核基因转录调节中的两种主要成分

在遗传信息传递过程中基因的转录是第一个具有高度选择性的环节,而转录起始作为基因表达调控的第一步就更具有关键性的意义,因为在这一步的调控对细胞的生化性质将具有极大的影响,既决定基因表达的类型,又关系到基因表达的水平。

#### (1) 顺式调节元件

顺式调节元件( cis-acting regulatory element)是指 DNA 序列上的一些对基因表达有调节活性的特定调控序列( regulatory sequence)。这种序列上分布着调节蛋白的结合位点( regulator site)。为使 RNA 聚合酶 II 以最大速率把基因转录为 RNA,许多顺式作用的调节元件必须协同作用。根据这些元件在 DNA 上的相对位置可以把它们大致分为 3 类:靠近转录起始位点的启动子,这是 RNA 聚合酶 II 结合的位置;近启动子元件( promotor proximal element),其序列上的一些位点在与调节蛋白结合后可反过来协助 RNA 聚合酶 II 结合到启动子上(图 15-1),以及增强子( enhance)和沉默子( silence)。不论是启动子、近启动子元件,还是远距离的增强子或沉默子,都是不同反式作用的 DNA 结合蛋白的靶位点。

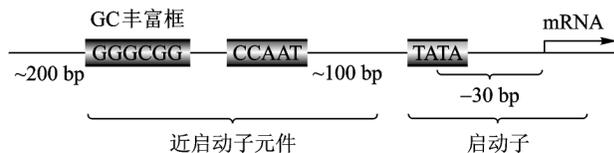


图 15-1 真核生物转录起始位点的上游序列(引自 Griffiths 等, 2005)

在许多真核生物中,仅有近距离的调节元件还常不足以引发高水平的转录。从启动子进行的转录往往还受远离起始点的调控元件的刺激。如增强子可大大增强同一 DNA 分子上启动子的转录效率,因此对转录起到激活和正调节作用。沉默子的作用则相反,结合了阻遏物的沉默子序列抑制激活因子并降低转录效率,甚至使整个调控系统失活。

#### (2) 转录调节蛋白

DNA 上的顺式作用元件不论是启动子还是增强子或沉默子,都需要与专一序列 DNA 结合蛋白( sequence specific DNA-binding protein)相结合,才能行使对转录及其相应过程的调节功能。于是,人们便把这样一些直接或间接地识别或结合各顺式作用元件 8~12 bp 核心序列,并参与调控靶基因转录效率的一组蛋白统称为基因转录调节的反式作用因子( trans-acting factor)。它们的功能不同,分别由 DNA 上不同的基因座编码。依据它们的功能大致可分为转录因子( transcription factor)、激活因子( activator)、辅激活因子( coactivator)和阻遏物( repressor)。

在真核生物中,3 种 RNA 聚合酶都不能直接与启动子接触,只有当相关的转录因子在启动子区构架了“平台”后,聚合物才能进入启动子区参与转录起始复合物的组装(详见第 3 章)。根据转录因子的功能,那些在转录起始时首先与核心启动子结合并组成转录起始基本装置( basal apparatus)的调节蛋白常统称为通用转录因子( general transcription factor, GTF),如 TFII D 等。它们的主要作用是将 RNA 聚合酶定位在启动子上。那些能识别特定的共有元件,结合到启动子或增强子短序列上,以加强启动子上基本装置的效能,并提高转录频率的调节蛋白就特称为激活因子。它们在真核基因转录

调控中占有十分重要的地位,是基因表达调控的核心成分之一。辅激活因子的特点是在激活因子和基本转录装置间提供一个“连接”,利用蛋白质与蛋白质之间的相互作用来辅助转录激活,而不是通过自身与 DNA 的结合来促进转录作用。阻遏物是近年来在真核细胞中发现的一类与转录起始相关的结合蛋白,它们结合于上游启动子元件或更远的沉默子位点,可抑制转录起始。

## 15.1.2 转录调节蛋白的结构和功能

### (1) 转录调节蛋白的结构

与近启动子或增强子、沉默子相互作用的转录调节蛋白,为行使其结合 DNA 和激活阻遏转录的功能,在结构上至少有两个功能域:一个结合顺式元件的 DNA 结合域和另一个与其他结合蛋白相结合以调节转录的激活域。

① 转录调节蛋白的 DNA 结合功能域 与原核生物类似,大多数真核生物的调节蛋白也是采用二聚体的方式与 DNA 结合,以  $\alpha$  螺旋插入大沟来识别特定的 DNA 序列中的碱基对。这种螺旋称为识别螺旋(recognition helix, HTH)或序列阅读螺旋(sequence reading helix)。大多数 DNA 结合结构域都是带正电荷的,因之能与带负电荷的 DNA 磷酸基相吸引。当调节蛋白插入到识别区域的 DNA 双螺旋大沟中后,便与其中的碱基对产生一系列的分子接触点,氨基酸和 DNA 碱基对的接触边缘形成氢键、离子键来调节相互间的作用。尽管蛋白质与 DNA 之间每个接触点的作用力都比较弱,但 20 个或更多的这种接触加在一起,就相当强,而且具有高度的专一性。事实上,蛋白质与 DNA 之间的相互作用是当前已知生物学中最为紧密、最为专一的一种分子间相互作用。真核生物调节蛋白的 DNA 结合功能域各具不同的结构域,目前,研究得较为清楚的有 4 种类型: $\alpha$  螺旋-转角- $\alpha$  螺旋结构域(helix-turn-helix motif),锌指结构域(zinc finger motif),亮氨酸拉链结构域(leucine zipper motif),螺旋-袢-螺旋结构域(helix-loop-helix motif)。

② 转录调节蛋白的激活域 激活域与 DNA 结合域不同,它没有特定的结构模式,而是以一种黏性表面与前起始复合物之间形成接触点来起作用。激活域是由许多重复的、活化能力较弱的小单元组成,每个单元是一段短的氨基酸序列。这样的小单元越多,激活域的活力也就越强。激活域是按其所含的基本氨基酸的成分归类,大致可分为以下 3 类。酸性功能域(acidic domain):富含酸性氨基酸(如天冬氨酸和谷氨酸),是最常见的一类激活功能域,如 GAL4。富含谷氨酰胺的功能域(glutamine-rich domain):多见于 DNA 结合域是同源异形域的转录因子,曾在哺乳动物激活因子 SP1 中发现。富含脯氨酸的功能域(proline-rich domain):只在哺乳动物激活因子 TF1 中发现。

③ 转录调节蛋白的模块式结构特征 有关酵母、果蝇、哺乳动物中的一些激活因子的研究已经证实,这些调节蛋白都是由结构上独立的 DNA 结合域和激活域共同组成的模块式结构。模块中含有一个或几个激活功能域,它们通过可弯曲的蛋白质区域连接到一个专一序列的 DNA 结合功能域上。DNA 结合域不仅履行其结合 DNA 的功能,有时还可对转录激活起一定的作用。激活功能域是通过与转录相关的其他蛋白质的结合来发挥功能。由于有把 DNA 结合域连接到激活域上的可弯曲区,这就使行为上独立的 DNA 结合域和转录激活域发生距离变动时,还能保持着它们之间的协同功能。甚至,当转录因子结合 DNA 的位置错移,其对应的激活域仍能彼此共同作用。

### (2) 转录调节蛋白的主要功能

激活因子可协助前起始复合物的组装,并整合信息,以增强转录效率。在真核基因转录中,聚合酶与启动子的结合较为复杂。基因转录的起始除需要一系列的通用转录因子外,还要有其他转录调节蛋白,如激活因子、中介蛋白复合体等的协助,才能使 RNA 聚合酶组合到转录装置上,并稳定地结合于启动子。在这一组合过程中,结合于 DNA 的激活因子与转录装置之间所发生的相互作用是借助于中介蛋白复合体(mediator complex)来完成的。中介蛋白由一些亚基组装而成,它实际上是一种辅激活因子,通过蛋白质-蛋白质之间的相互作用,在激活因子与基础装置间起到“桥梁”的作用。中介蛋白表面有一个部位与聚合酶的羧基末端域(carboxy terminal domain, CTD)结合,另一个部位则与

激活因子相互作用。于是激活因子通过中介蛋白把聚合酶引导到基因上。除 RNA 聚合酶外,激活因子还可与转录装置中其他复合物(如 TF II D)相互作用,把它们引领到基因上(图 15-2)。正是通过上述这些作用,激活因子促进完整的前起始复合物的形成。

人类  $\beta$  干扰素是抵御病毒感染的一种主要的蛋白质,对其基因中增强子的分析,可为转录调节蛋白间的相互作用提供很好的例证。 $\beta$  干扰素的增强子中有 4 个控制元件,它们可在 DNA 双螺旋的同一面同时结合 4 种不同的转录调节蛋白,在增强子上产生一个多蛋白复合物,称为增强体(enhanceosome)(图 15-3)。增强体一旦形成,其中的激活因子将同转录装置联系,并高水平地激活基因转录。增强子中各元件上分别结合的激活因子具有高度的协同性,保证了信号的紧密结合。这些激活因子结合的协同作用是由一种非组蛋白 HMG I (high mobility group I) 促进的, HMG I 结合于 DNA 螺旋的小沟中,在使 DNA 弯曲形成增强体和激活因子之间的协同作用中起着关键的作用。

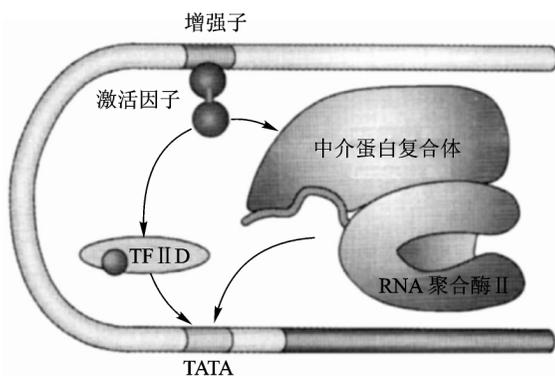


图 15-2 真核细胞中激活因子引领转录装置以激活转录起始(引自沃森, 2004)

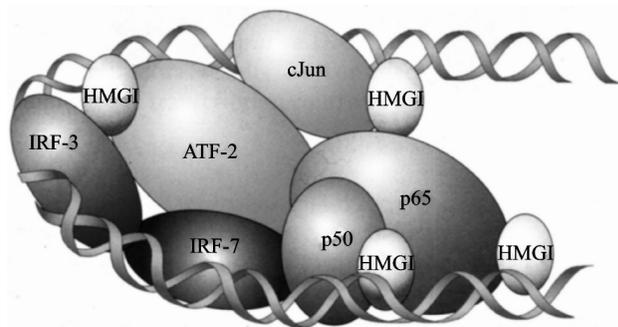


图 15-3 人类中在  $\beta$  干扰素增强子上形成的增强体(引自 Lodish 等, 2004)

IRF-3 和 IRF-7 是两个单体转录因子, Jun/ATF-2 和 p50/p65 (NF- $\kappa$ B) 是两个异二聚体转录因子, 它们结合到增强子的 4 个不同的控制元件上, 是在 HMG I 的促进下协同结合形成增强体

**阻遏物对基因表达的影响** 原核生物中,阻遏是一种较普遍的调节方式,并已分离出多种阻遏物,如乳糖操纵子的 Lac I、色氨酸操纵子的 Trp R 等。它们多是通过与启动子重叠位点的结合,或结合于启动子的邻近区,来阻断 RNA 聚合酶与启动子结合,或与结合在启动子上的聚合酶相互作用,或干扰激活因子的功能等,从而抑制基因转录的起始。在真核生物基因表达方面,过去人们多侧重于激活因子对基因转录的作用,但近年已愈来愈多地发现了抑制转录起始的调节蛋白——阻遏物的作用,它们结合于上游近启动元件或更远的沉默子中的位点,阻遏基因表达。不同的阻遏物以不同的方式行使其抑制功能。如有些阻遏物是通过与激活因子竞争同一增强子元件来阻扰激活因子对同一序列的接近,以致无法激活转录[图 15-4 a]。

有的阻遏物则是采取直接抑制激活蛋白的方式,迫使后者无法发挥作用。这类抑制有两种:一种是阻遏物结合到激活因子的 DNA 结合域上,使激活因子不能与增强子结合;另一种则是阻遏物结合于激活因子的激活域上,这种情况可见于酵母的半乳糖系统。这种被封锁的激活因子虽然还能与增强子结合,但却不能执行其激活功能[图 15-4 b]。

上述有关阻遏物的“竞争”和“压制”导致转录因子激活作用降低,但对基础转录没有影响。只是当一些真核细胞的阻遏物结合到近启动子相邻位置,阻止聚合酶接近启动子,这就直接消除了一切转

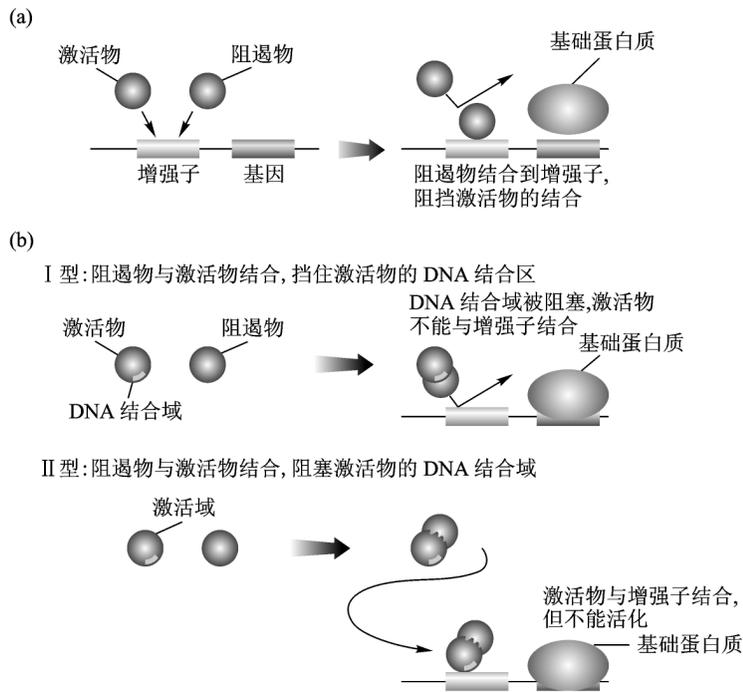


图 15-4 阻遏物抑制基因表达(引自 Hartwell等, 2004)

(a) 阻遏物与激活物竞争结合 (b) 遏制

录活动;或者是阻遏物结合到启动子上游的 DNA 位置。由于增强子与启动子间的袢环,使得该阻遏物得以与转录因子基本装置接触,从而降低其基础转录水平。

### 15.1.3 染色质修饰与基因表达

#### (1) 核小体重塑与转录起始

在真核生物中,由于 DNA 是包装到核心组蛋白上,根据机体发育的需要,基因活性时常发生改变,因而染色质构型的状态对基因表达有着重要的影响。DNA 酶 I 敏感性的实验已经证实,转录的 DNA 比处于关闭状态的更容易受核酸酶的攻击。那么,转录的 DNA 是否仍包装在核小体上? 转录因子、RNA 聚合酶又如何与 DNA 接触,而进行基因转录? 遗传学和生物化学的研究表明,基因活跃转录的常染色质中,其核小体位置有所变动,组蛋白八聚体与 DNA 的结合是处于动态变化之中。也就是说,转录时, DNA 区段可从紧密相互作用的组蛋白八聚体上暂时解离出来。现已证实,基因转录时,染色质中核小体的动态变化受到多蛋白复合物的修饰调节。

核小体的修饰有两种形式:一种是增加组蛋白肽链末端的化学基团,如乙酰转移酶可增加组蛋白 N 端的乙酰基团,可以激活染色质内那些难以接触到的基因;而组蛋白 N 端某些部位的甲基化则可抑制基因转录。另一种则是重塑核小体,修饰物是利用 ATP 水解释放的能量,使核小体组蛋白核心改变位置,暂时脱开 DNA,或是使核小体核心沿 DNA 滑动,促进高度有序的染色质结构松开。这种在一定能量下核小体移动或改组的过程称为染色质重塑(chromatin remodeling)。而那些有助于核小体移动的蛋白质复合物便称为核小体重塑复合物(nucleosome remodeling complex)或染色质重塑复合物(chromatin remodeling complex)。

目前研究得最为深入的染色质重塑复合物是在面包酵母中发现的 SWI/SNF[SWI/SNF 的命名起因于两种突变——开关抑制和蔗糖培养基上不发酵(switching inhibition and sucrose nonfermenter)]。SWI/SNF 约由 8 个蛋白质组成,它能使组蛋白八聚体沿 DNA 分子侧移,把覆盖于 TATA 序列的核小体移开,使转录因子和聚合酶 II 能够与 DNA 接触,促进基因表达(图 15-5)。

现在从多种生物中已分离出 SWI/SNF 的类似物,表明在任何一个细胞中都有多种类型的核小体重塑复合体,进一步证明这类蛋白质是真核基因调控制体中不可缺少的部分。

## (2) 组蛋白修饰与组蛋白密码假说

① 组蛋白乙酰化和脱乙酰化与基因表达 一个基因是否表达与其启动子所在位置以及周围的染色质结构状态密切相关。一般说来,活化的基因处于染色质伸展状态。可见,转录前染色质结构的变化是转录的前提条件。染色质结构通过某些修饰而发生相应的改变,如核小体核心组蛋白上某些氨基酸就可被共价修饰。修饰作用有多种类型,如乙酰化、甲基化、磷酸化、泛素化等。其中,组蛋白 N 端的乙酰化修饰与基因表达的增强相关,直接影响核小体的结构,使组蛋白八聚体与 DNA 的结合松动,有利于基因转录。体外实验显示,组蛋白末端的乙酰化能够使 DNA 更容易受到 DNA 酶的攻击,也更容易与转录因子结合。其机制可能是被乙酰化的 N 端携带的正电荷减少,以致组蛋白八聚体与 DNA 结合的稳定性降低,而且也降低了核小体排列的紧密程度,为转录因子、RNA 聚合酶等与转录相关的蛋白质因子结合到启动子附近的 DNA 序列,发挥调节作用提供了可能。

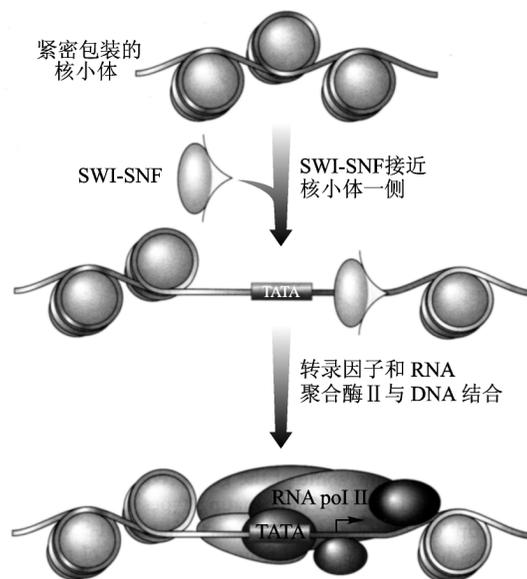


图 15-5 染色质重塑(引自 Griffiths 2005)  
展示 SWI/SNF 使 TATA 框暴露,便于 RNA 聚合酶复合物的结合

组蛋白乙酰化是由组蛋白乙酰基转移酶(histone acetyltransferase, HAT)来介导的。真核生物的 HAT 酶有两群:A 群与转录有关,B 群与核小体组装有关。大量研究证明,基因表达的活性与启动子序列附近的组蛋白 N 端乙酰化和 HAT 复合体在这个区域的聚集程度呈正相关性。那么,组蛋白乙酰化究竟是怎样使染色质结构发生改变的呢?

有关酵母的遗传和生化研究,发现有一个大的多蛋白复合物,其中包含的酵母调节蛋白 Gcn5 本身就具有乙酰化酶活性,而且 Gcn5 还是一些增强子与其靶启动子间连接复合物的组成成分之一。基因转录时,酵母激活因子 Gcn4 的 DNA 结合域结合于它所调节的基因的专一上游激活序列(UAS),而其酸性激活域与一个包含有 Gcn5 催化亚基的多蛋白组蛋白乙酰化酶复合物相互作用。随后 Gcn4 结合位置附近核小体的组蛋白 N 端被高度乙酰化,染色质解缠绕,起始所需要的通用转录因子得以结合(图 15-6)。由于组蛋白 N 端是染色质 30 nm 纤维形成所不可或缺的,被乙酰化修饰之后,便影响了组蛋白进一步形成高级染色质结构的能力。现已观察到,这种被修饰的核小体几乎不参与 30 nm 染色质纤维的形成。

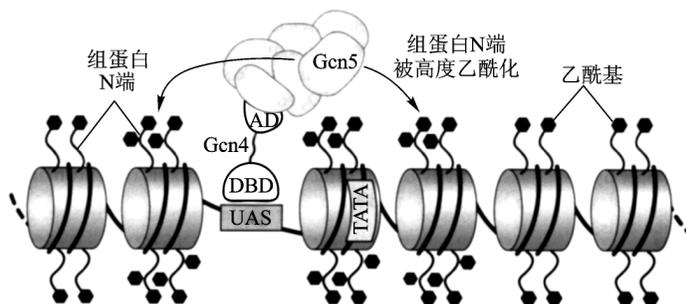


图 15-6 组蛋白 N 端高度乙酰化机制示意图(引自 Lodish 等 2004)

真核生物中也有类似于酵母的这种激活机制。例如哺乳动物中的多功能域蛋白 CBP 和 P300, 它们功能相似, 且密切相关, 常被看作是一个单一激活类型。P300/CBP 也具有 HAT 活性, 它实际上是一个辅激活因子, 可将激活因子和基础转录装置联系起来。它能与各种激活因子相互作用, 包括激素受体、AP-1、MyoD 等。另外, 最大的 TFII D 亚基也具有组蛋白乙酰酶活性, 通过对 TATA 框附近的核小体组蛋白 N 端尾部乙酰化而起到一个辅激活物的作用。

在组蛋白尾的修饰中, 常有一些调节蛋白的参与, 它们与所修饰的氨基酸特异结合。这类调节蛋白都具有一个保守的蛋白质结合域。如与组蛋白 N 端乙酰化的赖氨酸结合的调节蛋白, 其结合域称为溴基域 (bromo domain) (图 15-7)。它们与乙酰化的组蛋白末端构成的复合体可进一步修饰乙酰化的区域, 使常染色质得以维持, 活化基因表达。

乙酰化作用是可逆的, 组蛋白 N 端氨基酸上的乙酰化基因可以被组蛋白脱乙酰基酶 (histone deacetylase, HDAC) 移走。许多研究也显示, 启动子附近组蛋白的脱乙酰基化和 HDAC 在该区域的聚集与基因表达受到抑制有关。在酵母中已发现 Rpd3、Sir3、Sir4 这样一些转录阻遏物, 它们具有 HDAC 活性, 当结合到基因调节区时, 由于使附近组蛋白发生脱乙酰基作用而抑制基因转录。

② 组蛋白甲基化与转录调控 组蛋白 N 端的甲基化发生在精氨酸 (R) 和赖氨酸 (K) 残基上。组蛋白甲基化的这些反应由不同的甲基化转移酶 (methyltransferase) 催化。从目前发现的情况看, 组蛋白上精氨酸的甲基化常伴随着转录的激活, 而赖氨酸残基的甲基化则因赖氨酸所在的位置不同而有差异。赖氨酸甲基化发生在组蛋白 H3 的第 4、9、27、36、79 (K4、K9、K27、K36、K79) 位及 H4 第 20 (K20) 位上。其中, 在酵母和哺乳动物细胞中 H3K4 和 H3K36 位点被甲基化可以激活转录; 而 H3K9、K27、K79 和 H4K20 的赖氨酸甲基化则可抑制转录。

位置效应花斑抑制子 [(suppressor of position effect variegation, SUV(var))] 是在果蝇中发现的第一个组蛋白赖氨酸甲基化转移酶, 它催化 H3-K9 的甲基化。在其他真核生物中也陆续发现类似的组蛋白甲基化转移酶, 如哺乳动物中的 SUV39H1, 酵母中的隐蔽位点调节子 (cryptic loci regulator, CLR4)。它们的催化位点都具有一个保守的 SET 结构域。在组蛋白尾甲基化修饰中, 甲基转移酶与含有保守的 chromo domain 的调节蛋白专一结合, 有助于维持核小体结构, 形成异染色质, 抑制基因表达。发生在 H3 球形核心 79 位赖氨酸 (K79) 上的甲基化对端粒异染色质形成可能至关重要。

关于组蛋白甲基化与异染色质形成的机制, 有学者提出相关的模型: 首先由脱乙酰酶解除组蛋白 H3 N 端第 9 及 14 位赖氨酸的乙酰化, 随后 SUV39H1 作用 H3 尾, 创造一个甲基化的定位信号, 于是调节蛋白——异染色质蛋白 1 (heterochromatin protein 1, HP1) 的 chromo 域与这个被甲基化的赖氨酸相互作用, 成为形成失活染色质的一个靶位。由于 HP1 分子自身的凝聚能力而彼此作用, 染色质失活区域得以延伸而形成异染色质区。在酵母中发现的沉默信息调节 (silent information regulator, SIR) 复合物 (其中 SIR2 具有甲基化转移酶活性), 它在异染色质形成中的作用, 与上述哺乳类的 HP1 有着类似的反应机制, 即这些调节蛋白在染色质中的接触位点都是组蛋白中被修饰的 N 端尾。

组蛋白甲基化除在基因转录表达方面起重要作用外, 其功能还体现在 X 染色体失活、基因印记 (gene imprinting) 中。因而组蛋白甲基化也是表观遗传 (epigenetic) 修饰的一种方式 (详见第 3 章)。

③ 组蛋白密码假说 结合在核小体中的组蛋白, 其 N 端尾从 DNA 缠绕的核小体中伸出, 在几

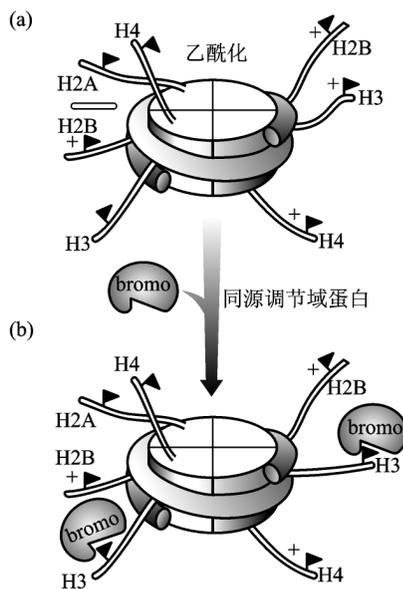


图 15-7 具有 bromo domain (溴基域) 的调节蛋白与组蛋白乙酰化末端的结合

个特定位置上的氨基酸能被各种修饰酶所修饰(图 15-8)。组蛋白尾上的这些修饰为其效应蛋白提供了结合位点。通过这些效应蛋白本身的作用,或是借助它们募集其他辅助蛋白(辅激活因子或辅阻遏物)的间接作用,来改变核小体的构象以及染色质的性质,从而进一步影响 DNA 的复制和基因表达的调控等。

由于不同的修饰酶在组蛋白尾上各具特定的靶位,说明其修饰的专一性(表 15-1)。这种修饰可用于调节基因的活性,因此有的学者认为,核心组蛋白 N 端上不同形式的修饰具有不同的作用,可以被阅读,故特称为组蛋白密码(histone code)。所谓组蛋白密码假说(histone code hypothesis)也就是指在组蛋白 N 端特定位置上的独特修饰组合,可以构成一个个“密码”,能影响与组蛋白-DNA 复合物相互作用的那些蛋白质,以及后续的基因调节状况。

### (3) DNA 甲基化与基因表达

① DNA 甲基化调节基因的转录 在染色质中,甲基化作用不仅发生在组蛋白上,而且也发生在 DNA 上。DNA 中大多数甲基化的位点为胞嘧啶,尤其是 CpG 岛(CpG island)。DNA 甲基化是由 DNA 甲基化酶(DNA methylase)催化,最常见的甲基化形式是把甲基基团加到胞嘧啶环的 5'位置上,形成 5'-甲基胞嘧啶(<sup>m</sup>CpG)。哺乳动物约有 5%的胞嘧啶被甲基化成为<sup>m</sup>CpG 的形式。DNA 的甲基化可调节基因转录活性。在原核生物中,虽有 A、C 两种甲基化核苷酸,但其转录活性相差几千倍。而在真核生物中,DNA 甲基化与否,转录活性的差别可达上百万倍。可见,甲基化的调节作用在真核生物中更为重要。另外,在一些酵母、果蝇等低等真核生物的基因组中没有明显的 DNA 甲基化,表明 DNA 甲基化是脊椎动物细胞中转录调节的重要环节。

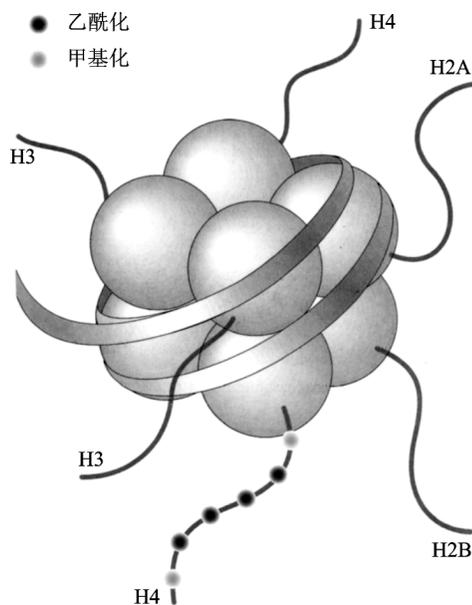


图 15-8 核小体组蛋白尾乙酰化和甲基化示意图(引自 Griffiths 等, 2005)

表 15-1 组蛋白修饰对染色质的结构和功能的影响

组蛋白	位置	修饰作用	功能
H3	Lys-4	甲基化	染色质凝缩 为 DNA 甲基化所必需
H3	Lys-9	甲基化	
H3	Lys-9	甲基化	
H3	Lys-9	乙酰化	
H3	Ser-10	磷酸化	阻止在 Lys-9 的甲基化 端粒沉默
H3	Lys-14	乙酰化	
H3	Lys-79	甲基化	
H4	Arg-3	甲基化	核小体组装 果蝇 X 染色体失活
H4	Lys-5	乙酰化	
H4	Lys-12	乙酰化	
H4	Lys-16	乙酰化	
H4	Lys-16	乙酰化	

注:在组蛋白中大多数修饰位置有一个信号、修饰专一类型,但有些位置具有一个以上的修饰类型,个别功能可能与一些其他修饰有关。

DNA 甲基化对基因表达的调节主要表现为抑制转录活性。在特异表达某些基因的组织中,活性基因附近的甲基化程度远低于 30%左右;而哺乳动物异染色质内的核 DNA 约有 80%的 CpG 被甲基化,说明甲基化程度与基因表达呈负相关性。DNA 甲基化抑制基因表达的机制目前还不明确,有一种可能是由于 DNA 甲基化直接抑制了转录因子的结合,不能形成转录复合物,从而也就抑制了基因转录活性。不过,这种情况并不普遍,因为很多与转录因子结合的 DNA 序列不存在甲基化的 CpG,而且

即使有些转录因子的结合序列含有甲基化 CpG,却并不影响转录因子的结合,如 Sp1 (specificity protein 1)与 DNA 的结合就不受甲基化的影响。另一种可能的机制是甲基化的 DNA 对一些专一阻遏物具有特异亲和力。目前已发现多种与甲基化 CpG 结合的蛋白质,如 5-甲基胞嘧啶结合蛋白(5-methylCpG binding protein, MeCP) MeCP 1 和 MeCP 2 它们与 CpG 甲基化 DNA 序列结合,阻遏基因转录。

MeCP 1 所结合的 DNA 序列常需要有 10 个以上的甲基化 CpG,这一蛋白广泛存在于许多组织中。MeCP 2 含有一个甲基化 DNA 结合区(methylCpG binding domain, MBD),因此在细胞内可以特异地结合在甲基化的 DNA 上。除 MBD 外,MeCP 2 还有一个转录抑制功能区(transcriptional repression domain, TRD)。推想,MeCP 2 的转录抑制机制是通过 MBD 与甲基化 DNA 结合,使抑制区 TRD 有机会与转录复合物或转录因子发生蛋白质与蛋白质间的相互作用,最终控制基因的表达。实验证明,MeCP 2 还可以与组蛋白脱乙酰酶结合,彼此作用,修饰染色质结构,控制转录活性。由此看来 DNA 甲基化与组蛋白甲基化很可能彼此相关。

当基因表达调控需要,甲基化的 DNA 可以发生去甲基化,那么去甲基化区域是如何建立并得以维持呢? 如果一个 DNA 位点未曾被甲基化,识别这种非甲基化序列的蛋白质可以保护它不受甲基化作用。一旦一个位点已经被甲基化,则可能有两种去甲基化方式。一种是与半保留复制相联系的被动方式,即在复制时阻断维持甲基化酶(maintenance methylase)的作用,因而在第二次复制周期之后,子代双螺旋即可去甲基化。另一种则是由 DNA 脱甲基酶直接催化去掉该位置的甲基化基团,或移走被甲基化的胞苷。近年虽已发现脱甲基化酶,但其作用机制还有待进一步研究。

② 甲基化与亲本印记 基因的转录水平与 DNA 甲基化的状况相关,具有活性的基因与失活基因比较,前者极少被甲基化。20 世纪 90 年代在哺乳动物中发现亲本印记(parental imprinting)的表现遗传现象。亲本印记是指一个基因的活性要依靠其亲本来源而决定,又可称为基因组印记(genomic imprinting)。在亲本印记中,某些常染色体基因由于 DNA 中的 CpG 双核苷酸中的 C 被甲基化而具有异常的遗传模式。如在小鼠中,编码类胰岛素生长因子(insulin like growth factor)的 Igf2 基因,由父本而非母本提供时表达。反之,另一个已知基因 H19 则是由母本而非父本提供时才表达。为什么同一细胞中来自于父母本染色体上的两对等位基因会出现这种相反的遗传现象? 原来这是因为在两个基因之间有一段印记控制区(imprinting control region, ICR),在父本染色体的等位基因上 ICR 被甲基化, H19 失活,而 Igf2 表达。在母本染色体上则是由 ICR 的绝缘子(insulator)功能来控制 Igf2。当 ICR 未甲基化时,可与 CTCF 蛋白(CCCTC-binding factor,一种序列特异性 DNA 结合蛋白,具有 11 个锌指结构)结合,从而阻止增强子激活 Igf2 的启动子(图 15-9)。

通常,在具有野生表型的个体中,往往不会注意到基因组印记的状况,因为其对应的基因座是有活性的,可以提供所需要的产物。但是如果一个印记基因的等位基因发生缺失突变,会造成什么后果

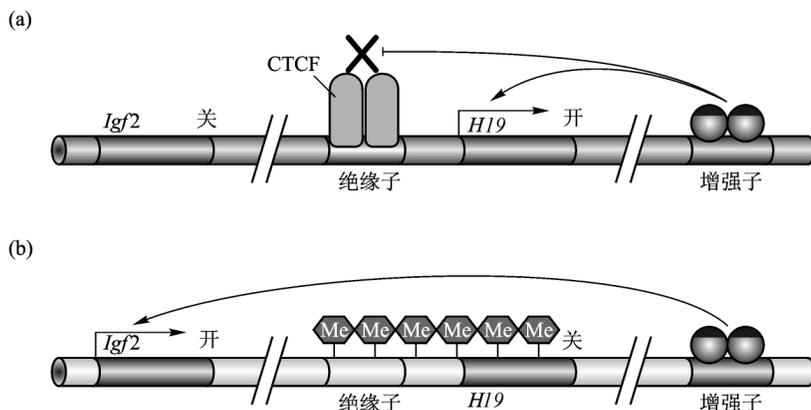


图 15-9 印记所控制的小鼠 Igf2 基因和 H19 引自沃森等, 2005

(a) 来自母方的染色体 (b) 来自父方的染色体

呢? 人类有些遗传病可以说明这种情况。如 Prader-W 综合征 (PWS) 和 Angelman 综合征 (AS)。前者症状表现为肥胖、手脚短小、性腺和外生殖器发育不全、智力低下; 后者症状是面容特殊、嘴大常呆笑、癫痫和严重智力低下 (也称快乐木偶症)。这两种综合征都与 15 号染色体 q11.13 小片段缺失有关, 因为在这一染色体区域内至少有两个核内小核糖核蛋白 N 多肽基因 (small nuclear ribonucleoprotein polypeptide N, SNRPN) 的拷贝, 它们具有不同的印记。如果一个基因是母亲印记, 那么, 当子代从父亲获得一条有缺失的 15 号染色体, 又从母亲获得一条无缺失但带有该基因印记拷贝 (失活的) 的 15 号染色体, 就将发生 PWS 综合征 (图 15-10)。相反, 如果另一基因印记是发生在父亲中, 子代的一对 15 号染色体中, 一条是来自母本的, 带有缺失; 来自父本的一条虽无缺失却具有失活的印记基因, 则将发生 AS 综合征。

“印记”基因并非编码在 DNA 序列中, 它是通过配子发生时 DNA 或染色质的某些表现变化来实行其效应。在哺乳动物中, 印记保持终生, 但随着原始生殖细胞的发育, 在配子形成前, 原有的

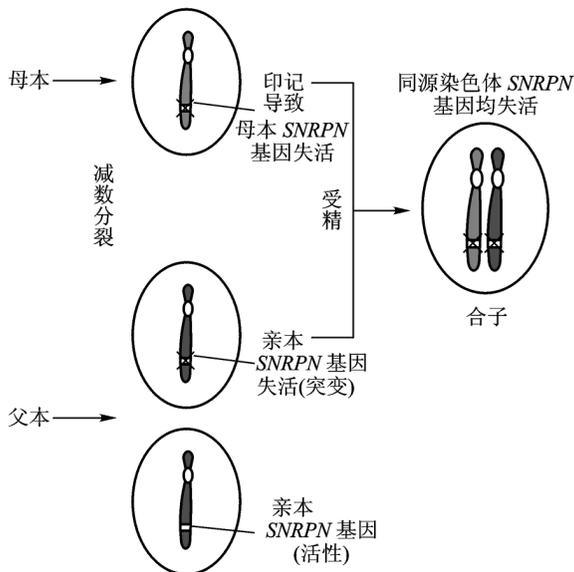


图 15-10 PWS 综合征的遗传起源 (仿自 Griffiths 等, 2005)

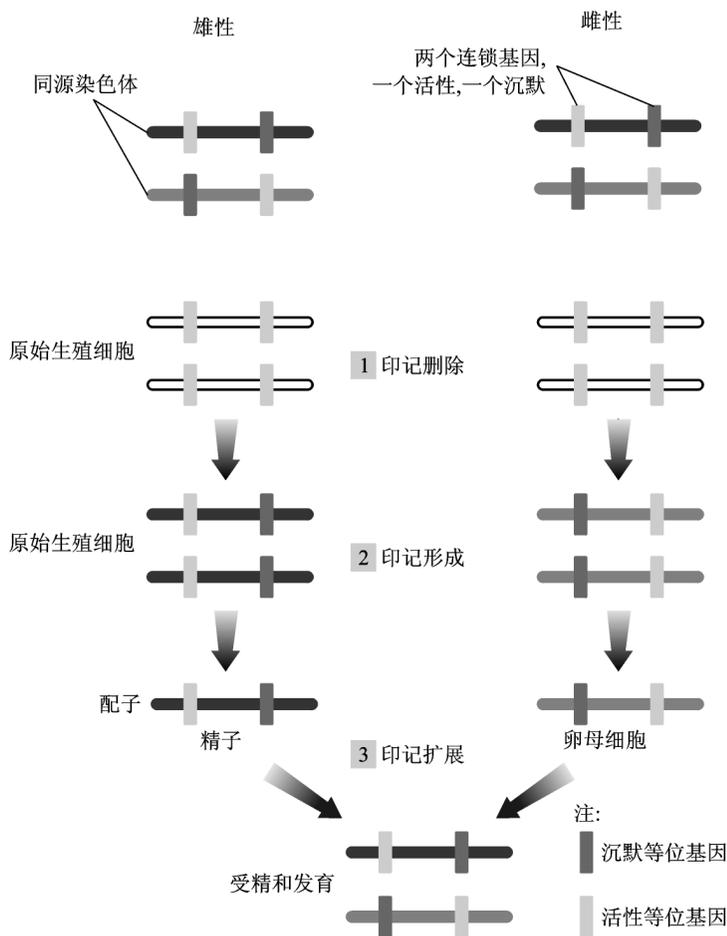


图 15-11 印记形成步骤

印记被消除,然后再按性别特异方式形成新的印记(图 15-11)。

### 15.1.4 基因表达的激素调节

多细胞真核生物的一些基因表达常受内、外激素(hormone)的调控。许多甾类激素如蜕皮素、皮质素、雌激素、睾酮、甲状腺素、糖皮质激素和一些多肽激素(如胰岛素)等,都可以促进某些基因的转录。甾类激素是一些较小的疏水性分子,可以穿过质膜进入靶细胞,与细胞质内或核内的相应受体形成复合物,导致受体分子三维结构甚至化学性质的变化。一般情况下,激素-受体复合物可直接进入细胞核,调控基因转录(图 15-12)。许多研究表明,激素通过同细胞内特定的受体蛋白分子相结合,激活特定的基因家族,能导致动物和人体在分化和发育上出现一系列复杂而又协调的反应。如昆虫变态时,蜕皮激素引起的多线染色体疏松区发生有规律的变化。用<sup>3</sup>H-尿嘧啶作为前体物进行的实验已经证明,这些多线染色体上疏松区的变化实际上就是基因正在转录的表现。

激素又是怎样来调控基因转录呢?在这方面,疏水性的甾类激素是研究得最为广泛和深入的。早在 20 世纪 70 年代,人们就发现激素可以使与它相结合的受体蛋白发生某些变化,从而结合到染色质上以促进转录。受体蛋白分子是一条由 800 个氨基酸组成的多肽链,含有 3 个功能区:羧基端激素结合区、中部 DNA 结合区和氨基端转录调控区。通常,在非激活状态时,受体上结合了一个含热激蛋白 90(Hsp90)的抑制蛋白复合物。一旦激素分子结合到受体的激素结合区,便导致

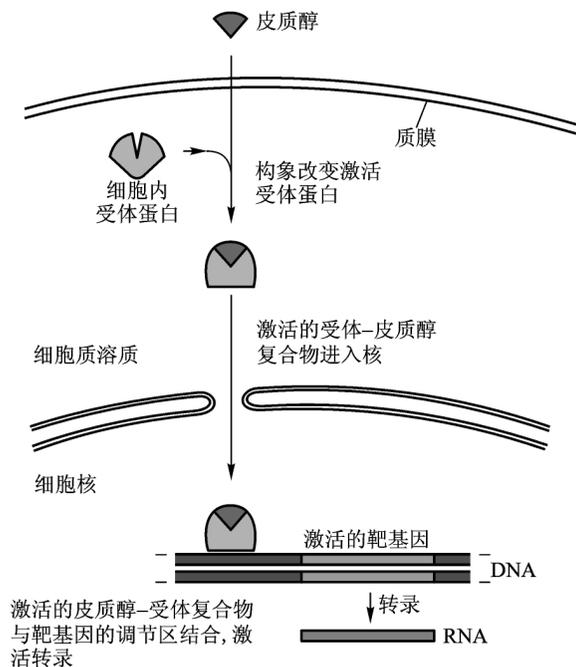


图 15-12 皮质激素激活基因调节蛋白

以激发基因转录(仿自 Alberts 等, 2004)

皮质醇扩散穿过质膜,与细胞质溶质中的受体结合。皮质醇-受体复合物经核孔进入核,皮质醇所结合的受体被激活,而后受体与 DNA 专一调节序列结合,激活基因转录

一旦激素分子结合到受体的激素结合区,便导致

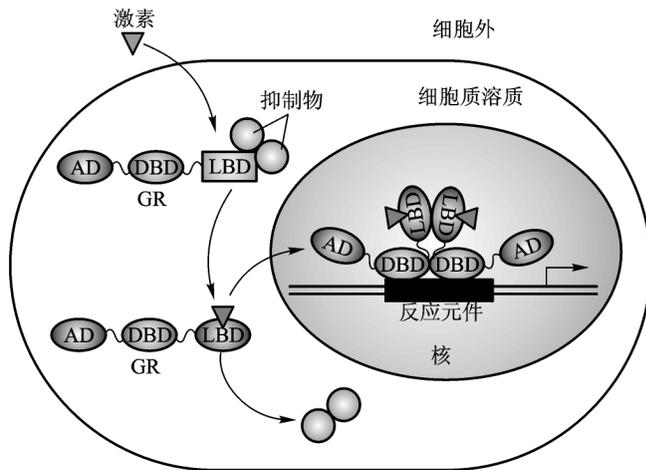


图 15-13 激素激活细胞内受体的机制(引自 Lodish 等, 2004)

无激素时,受体的配体结合域(LBD)保持与抑制蛋白相互作用,处于失活状态。激素进入细胞后便结合到受体的配体结合域,使受体发生构象改变,与抑制物解离。随后结合着激素的受体转移入核内,其暴露出的 DNA 结合域(DBD)便结合到 DNA 上,使配体结合域和 N 端的另一个激活域(AD)刺激靶基因的转录。GR 为糖皮质激素受体

上述抑制蛋白复合物解离,暴露出 DNA 结合区,受体被激活(图 15-13)。

从甾类激素的研究中已知,每一类激素有不同的受体蛋白,每个受体可作用于一套不同的调控位点,从而调节一套不同的基因。对一些甾类激素受体蛋白的比较研究,发现它们的大小虽不划一(如盐皮质激素受体长达 984个氨基酸,而维生素 D 只有 427个氨基酸),但都含有一个同源性很高的 DNA 结合域和一个同源性稍差的激素结合域,只是后者的氨基端差别大。另外,受体基因的结构也具有相似的高度保守性。因此人们把这些激素受体归为同一个大家族。这种激素受体复合物可和增强子或其他的基因调节序列结合,大大促进了基因的转录速率。受体蛋白的 DNA 结合区结构中有两个锌指结构,使受体能以二聚体形式更好地与 DNA 作用,调控基因转录。

## 15.2 真核生物基因转录后水平的调控

### 15.2.1 选择性剪接

基因转录合成前体 mRNA 后调节的第一步就是对此转录本的加工剪接。除常规的恒定剪接外,还可以进行其他的剪接修饰,如选择性剪接、反式剪接等。选择性剪接(alternative splicing)是指同一种 hnRNA 由于加工不同,而产生出不同的 mRNA 和蛋白质。这也就是说,来自一个基因的 RNA 前体分子不只包含一个内含子时,剪接发生在某个内含子 5'的供点和另一个内含子 3'受点,从而将两个内含子以及它们之间的全部外显子和内含子都剪接掉。例如,原肌球蛋白基因可以为脑、肝、骨骼肌、平滑肌和成纤维细胞编码  $\alpha$ 原肌球蛋白,由这一基因转录出的核前体 RNA 含有 11个外显子,通过选择性剪接,产生了上述不同细胞中的  $\alpha$ 原肌球蛋白(图 15-14)。

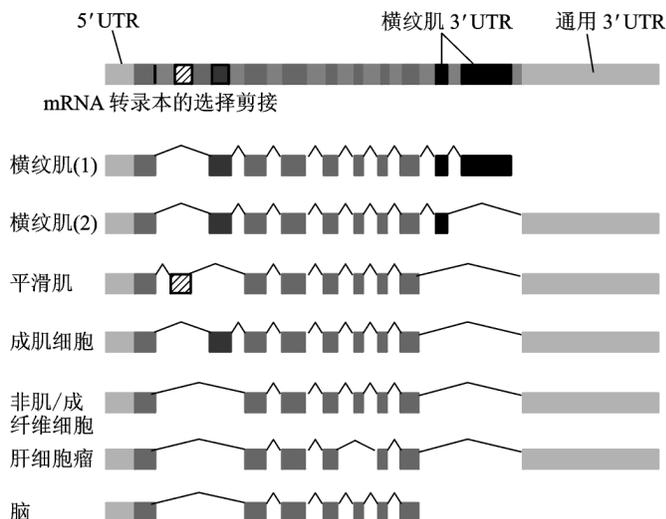


图 15-14  $\alpha$ 原肌球蛋白 hnRNA 转录本的不同剪接方式(自 Gilbert,2006)

同一个 hnRNA 转录本通过不同的剪接方式,能产生多种具有组织专一性的表达产物

由此可见,由一个基因所合成的 RNA 前体因选择性剪接而产生多种 mRNA,翻译出不同的蛋白质。这样就形成了一个相关的蛋白质家族,它们可以在不同发育阶段,不同组织或在细胞内不同的亚细胞结构中出现并发挥其功能。选择性剪接是一种重要的调节手段,使得一个基因所携带的遗传信

息在转录后有所扩展。

关于选择性剪接的调控方式,与常规调节方式基本相似,只是它是使用另外一套 snRNA 系统,并可受某些特定蛋白质的调控。

### 15.2.2 反式剪接

不论是常规的前体 mRNA 的修饰性剪接,还是可增加蛋白质种类的选择性剪接,一般都是在同一个前体 mRNA 上经过有序地删除内含子,把外显子连接起来形成一个成熟的可供翻译的转录物,这种剪接方式是顺式剪接(cis-splicing)。还有另外一种反式剪接(trans-splicing)方式,是把分别处在两条甚至 3 条前体 mRNA 上的外显子经过剪接,连接成为一条成熟 mRNA。通常,经过剪接在成熟的 mRNA 上游非编码区的 5'端拼接上一段剪接前导序列(splicing leader SL)或小外显子(mini exon)的 RNA 片段。这些片段原本不存在于相应的编码基因,而是由其他 DNA 链转录而来。

SL RNA 一般富含 AU,在种属间的 SL 其最大的同源部分只有 5 个核苷酸(AGUUU),而且在不同种属的成熟 RNA 中 SL 的比例和长度各不相同,如线虫中只有 10%~15% mRNA 的 SL 含有 22 个核苷酸,锥虫各种属 mRNA 的 SL 则有 39 个核苷酸。

目前知道锥虫只有反式剪接,其蛋白质编码序列中有许多串联的多顺反子,有非编码的间隔区,反式剪接可任意选择其中某一顺反子作为 3'受点,经剪接连成成熟 mRNA。这也就是说没有顺式剪接时,反式剪接体的形成可促使 mRNA 的成熟,并可有选择地通过核孔。反式剪接这样一种特殊的剪接方式也具有多种形式,可以用于克服基因所携带信息量的限制,提高遗传信息的使用率。现在已知,锥虫、线虫以及植物叶绿体与线粒体中,反式剪接是 RNA 的主要剪接方式。

关于反式剪接的机制尚不很清楚。一般认为也是在剪接体中进行。在 mRNA 前体 5'端,翻译起始点前 30~70 bp 处有通用剪接受点,与 SL 内含子的 3'端形成 2',5'-磷酸二酯键,于是前体 mRNA 的 5'端就成为 Y 型分支结构的右上支,随后以类似套索结构的方式切除分支,使两个前体分子的内含子以及 Y 型结构切除。

### 15.2.3 RNA 编辑

RNA 编辑(RNA editing)也是改变 mRNA 序列的一种方式,是指对转录后的 mRNA 编码区进行碱基插入、删除或替换,以改变来源于 DNA 模板的遗传信息,翻译出不同于基因原编码的氨基酸序列的蛋白质。1986年, R Borne 等首次发现原生动锥虫细胞色素 c 氧化酶亚基 II(Co II)基因转录产物 5'端有几个非基因编码的尿苷酸,而这些额外尿苷酸的出现正好形成完整的开放阅读框,校正了基因分子内的移码突变。后来在一些植物线粒体的 mRNA、某些病毒的 RNA 中,还有哺乳动物载脂蛋白 B(apolipoprotein B, ApoB)也都陆续发现了编辑现象。看来, RNA 编辑的结果不仅扩大了遗传信息,还可能是生物适应中的一种保护措施。

在 RNA 编辑中非常典型的一个例子就是哺乳动物载脂蛋白 B 的编辑。载脂蛋白在血浆脂蛋白的装配、运转和代谢中均有重要的作用。按相对分子质量的不同可分为肝型 ApoB(相对分子质量:  $5.12 \times 10^5$ )和肠型 ApoB(相对分子质量:  $2.41 \times 10^5$ ),但两者却是由同一基因编码。在有机体个体发育的早期,不论是肝还是肠组织中都只有最初的基因转录产物——ApoB 前体 mRNA。哺乳动物 ApoB 基因有几个外显子,其中有一个外显子中特定的 CAA 密码子是 RNA 编辑的靶子。随着消化功能的发展,小肠 ApoB mRNA 第 666 位的胞嘧啶在胞苷脱氨酶的介导下,进行位点特异性的脱氨基反应,由尿嘧啶替换胞嘧啶,使密码子由 CAA→UAA,翻译终止。在肝中 CAA 密码子没有进行编辑,翻译成谷氨酰胺,所以其 ApoB-100 有全长蛋白质(约 4500 个氨基酸),在小肠里的 ApoB-48 因编辑而成了截短的多肽(约 2100 个氨基酸)(图 15-15)。

其他酶促脱氨的 RNA 编辑还有腺嘌呤(A)的脱氨基作用。在此反应中由于腺苷酶的催化, A→I 即产生了次黄嘌呤。由于次黄嘌呤可以与胞嘧啶配对,所以这种编辑形式很容易改变 mRNA 编辑

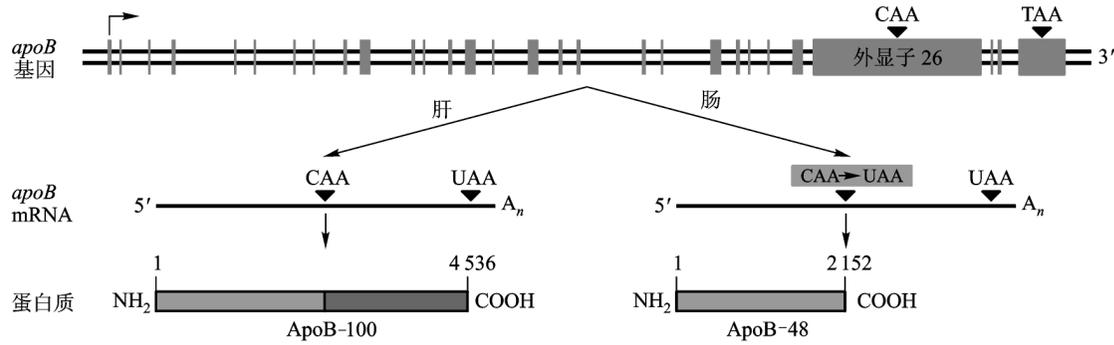


图 15-15 ApoB前体 mRNA 的 RNA 编辑(仿自 Lodish等, 2004)

肝中产生的 ApoB mRNA 与初级转录本具有一样的序列,其 mRNA 翻译成 ApoB-100。ApoB-100 含有两个功能域:一个是与脂质结合的 N 端域;另一个是与质膜上的 LDL 受体结合的 C 端域。小肠产生的 ApoB mRNA 中,外显子 26 的 CAA 密码子被编辑成 UAA 终止密码子。结果小肠细胞产生 ApoB-48 相当于 ApoB-100 的 N 端域

的蛋白质序列。哺乳动物大脑中所表达的一种离子通道蛋白就是通过这种方式编辑的。这是由于编码该蛋白的 mRNA 中段发生了一个核苷的替换,导致这种离子通道蛋白的一个氨基酸改变,即由精氨酸替换了谷氨酰胺,从而明显改变了该离子通道的  $Ca^{2+}$  通透性。如果不发生编辑,大脑的发育将严重受损。

在锥虫线粒体 mRNA 的研究中,发现了一种截然不同的 RNA 编辑方式。mRNA 转录后,在 mRNA 前体的特定位置上掺入了多个尿嘧啶(有时可能是删除几个尿嘧啶);而且有时插入数量非常大,甚至可达成熟 mRNA 核苷酸总数的一半左右。多个尿苷酸的掺入使 mRNA 的密码子和可读框发生改变,以致完全改变 mRNA 的原义。

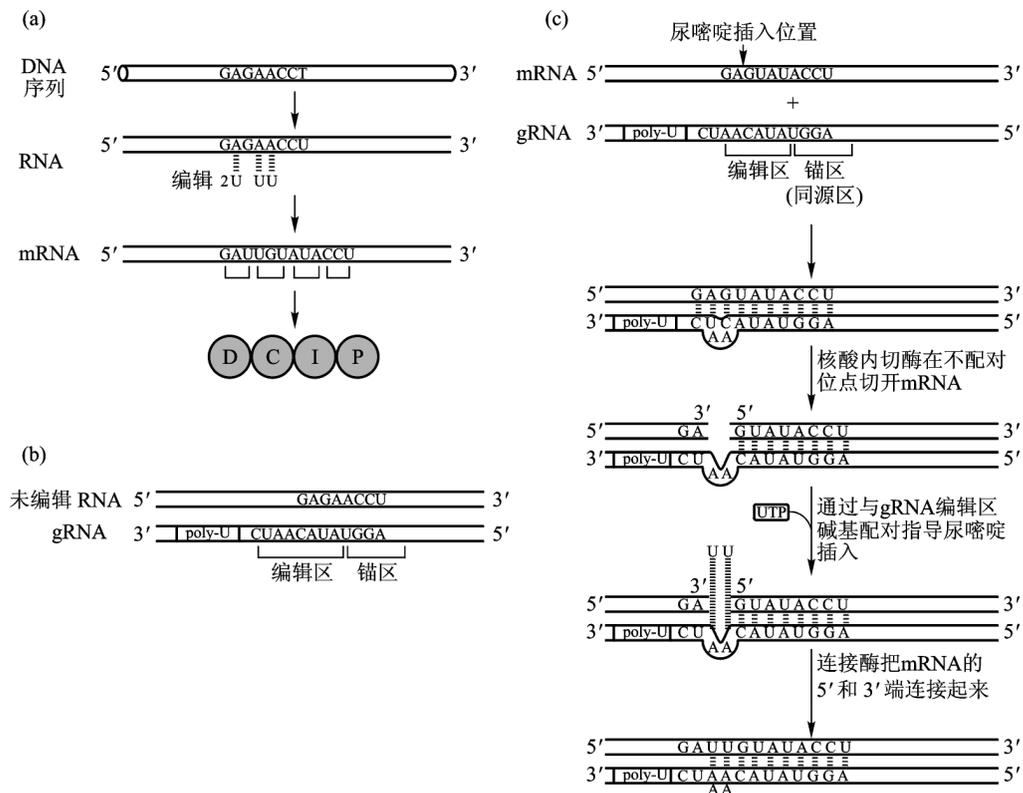


图 15-16 由 gRNA 介导的 RNA 编辑(自 Watson 等, 2004)

这些额外碱基是怎样掺入的呢? 进一步探索,发现有一类小分子 RNA 长 40~80 nt 是由不同于编码靶 mRNA 的基因编码。由于这种小分子 RNA 能引导额外的碱基掺入(或丢失)到 mRNA 中去,便被定义为指导 RNA(guide RNA, gRNA)。每种 gRNA 都可以分为 3 个区:第一区位于 5'端,称为锚区,负责指引 gRNA 到达它所编辑的 mRNA;第二区是用以精确定位所编辑的序列中碱基将要掺入的位置;位于 3'端的第三区是一段多聚尿嘧啶序列,可为所编辑的 mRNA 增添 U 或接受由 mRNA 上去除的 U 核苷酸。gRNA 编辑的方向是由 3'→5'进行,正好与转录或翻译的 5'→3'方向相反(图 15-16)。

## 15.3 真核生物基因翻译水平调控

### 15.3.1 mRNA 的稳定性

mRNA 翻译水平的调控主要是在翻译过程的起始阶段,其中包括两个水平上的调控:一是全局调控(global regulation),这种调控主要涉及蛋白质合成数量的整体变化,对所有 mRNA 的翻译都有影响,如翻译因子的可逆磷酸化。二是转录物专一性调控(transcript specific regulation),这种机制只作用于单个转录物或一小群编码相关蛋白的转录物,如哺乳动物中铁蛋白 mRNA 的调节。

一般,一种特定蛋白质其合成的速率同细胞质内编码它的 mRNA 的水平 and 稳定性成正比。例如真核细胞中有一些“持家基因”,其转录的 mRNA 一般均较稳定,这是因为由它们翻译而来的蛋白质是维持细胞基本功能所必需的。另外,高等真核生物一些高度分化的细胞中,其 mRNA 也十分稳定,如网织红细胞中已没有 RNA 的合成,可是血红蛋白合成速率却很高,表明血红蛋白的 mRNA 极为稳定。有时,由于终端分化的细胞中 mRNA 极其稳定,再加上 DNA 的扩增和强启动子的转录,真核细胞中有时还会出现令人惊讶的某种蛋白质的大量合成。如家蚕丝心蛋白的合成,其基因为单拷贝,带有强启动子,一旦开始转录,在短短几天内便可合成  $10^6$  个丝心蛋白的 mRNA 分子,每个 mRNA 分子又可以重复翻译产生  $10^5$  个丝心蛋白,可见其 mRNA 稳定性之强。但是,有一些蛋白质,特别是那些决定细胞周期的蛋白质,其合成必须按照严格的顺序开启和关闭,由这些基因转录的 mRNA 必然要相应更替。可见,mRNA 降解速率是其稳定性的衡量标志,降解速率慢,显示 mRNA 稳定。因而各种 mRNA 之间降解速率的差别也是真核基因在翻译水平上的调控内容之一。

mRNA 的正常翻译总是需要 5'端与 3'端之间的相互作用,终止翻译则需要破坏这二者间的正常作用。根据对哺乳动物细胞 mRNA 降解的种种分析表明,它们的降解过程都起始于 3'末端,或是去除尾部腺苷酸,或是在 3'-UTR 区发生专一性剪接,或是破坏并消除 3'端茎环结构。当 3'端 poly(A) 发生变化时,也影响到 5'端帽子结构,从而引起 mRNA 由 5'端向 3'端的降解(图 15-17)。

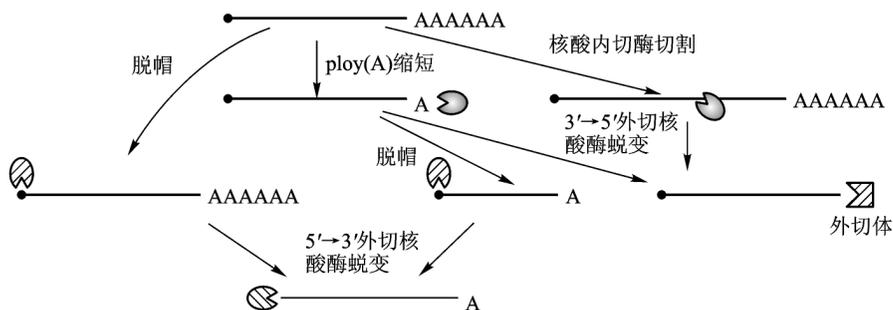


图 15-17 真核生物 mRNA 降解的途径

### 15.3.2 mRNA 非翻译区与翻译调控的关系

真核 mRNA 分子的非翻译区(untranslated region, UTR)既包括 5'端的帽子结构和 3'端的 poly(A)尾,也包括在 5'和 3'端的其他非编码序列。已知蛋白质的生物合成不仅与其 mRNA 的编码序列有关,而且还受到 5'端和 3'端非翻译区结构的调控。

#### (1) 5' 非翻译区与翻译调控的关系

翻译起始时,起始因子对“帽子”的识别非常重要。一般来说,尽管未甲基化的帽结构就可以保护转录本不会受 5'外切酶的降解,但还有许多研究进一步表明,只有当此帽被甲基化形成 m<sup>7</sup>G 状态时, mRNA 的翻译才更为有效。5'非翻译区除“帽子”外,其起始密码 AUG 所在位置旁侧序列的状况,先导序列的长度,以及 5'-UTR 本身的结构等也都对 mRNA 的翻译有不同的影响。

按翻译起始的扫描模型,以真核生物 mRNA 为模板的蛋白质生物合成是起始于最靠近其 5'端的第一个起始密码 AUG。但是在高等真核生物的细胞中,有些原癌基因和生长调节因子基因所产生的 mRNA,其 5'非翻译区内常有一个以上的 AUG,翻译起始并不一定遵循前述第一 AUG 规律。作为起始密码 AUG 与其旁侧序列关系密切。尽管有时邻近起始密码的序列并非最佳起始所需的共有序列,但是某些核糖体的翻译还总是开始于第一个位置的 AUG。对于大多数核糖体来说,只要第一个 AUG 的-3位为 A,就基本满足了由这一 AUG 来起始肽链合成的要求。这足以说明 AUG 位置的重要性。AUG 旁侧序列对于翻译起始的效率也很重要。Kozak 对 699 种脊椎动物 mRNA 翻译起始密码 5'和 3'两侧的核苷酸序列进行了分析,并提出了有关的共有序列(表 15-2),即起始密码 AUG 两侧核苷酸序列在-3位的 A 和+4位的 G 对于识别 AUG 具有最为显著的促进作用。如果-3位不是 A,则+4位的 G 是有效翻译起始作用所不可缺少的。1991年, D. R. Cavener 等也报道了对真核 mRNA 起始位点旁侧序列的统计学分析,进一步证实了-3位的 A 是真核 mRNA 的普遍规律。

表 15-2 真核 mRNA 起始密码 AUG 旁侧的共有序列

分类	共有序列				
	-3	-2	-1	+1	+4
脊椎动物	GCCGC	A G	CC	A UG	G
植物	A	A	CA	A UG	G C
原生动	AAAA UUUU	AAA UUU	A AA	A UG	A N
酵母	A U	A	A U	A UG	U CU

在 5'-UTR 中由第一个 AUG 至 5'帽子间的长度称为先导序列长度(leading length),它也会影响起始效率和翻译起始的准确性。Kozak 曾构建了一系列不同长度的 5'-UTR 转录物,实验结果表明,当第一个 AUG 密码至 5'帽结构的距离在 12 个核苷酸以内时,虽然起始密码是处于理想的旁侧序列之中,但却仍有一半以上的核糖体 40S 亚基会滑过第一个 AUG,而从其下游的 AUG 起始翻译。可见,5'-UTR 中第一个 AUG 密码距 5'帽子的位置太近时,不会被 40S 亚基识别。当 5'-UTR 长度处于 17~18 个核苷酸之间,此时体外翻译效率与其长度成正比。体内的实验也证实,适当加长这段距离还可减弱 AUG 5'端二级结构对翻译的抑制作用。

至于 5'-UTR 二级结构对 mRNA 翻译的影响,这是因为 5'-UTR 中有时存在碱基配对区。可以形成发夹式或茎环二级结构阻止核糖体 40S 亚基的迁移,对翻译起始有顺式阻抑作用。一般,这种碱基配对区愈长或 G+C 的含量愈高,发夹结构就愈稳定,其抑制作用便愈强。

总的看来,一个长度适当、没有高级结构,且上游又具有合适起始密码 AUG 的 5' 非翻译区,对于 mRNA 的有效翻译是必要的,而一个 G+C 含量高和富含 AUG 的复杂 5'-UTR,不论其长度如何则都将有碍于翻译的起始作用。

### (2) 3' 非翻译区与翻译的调控

真核 mRNA 的 3'-UTR 包括终止密码、poly(A) 尾以及前二者间的非编码序列,它们在翻译过程中同样具有重要的调控作用。真核生物 mRNA 翻译中 3 个终止密码的使用情况不同:UGA 在脊椎动物和单子叶植物中的使用频率最高;UAA 是其他真核生物中最主要的终止密码;而 UAG 的使用频率最低。对终止密码旁侧序列相对 GC 含量的分析并与 5' 非翻译区进行比较之后,发现终止密码的选用在很大程度上受 mRNA 中 GC 含量的影响。不同种类 mRNA 中,紧邻终止密码 3' 端的核苷酸在分布上具有一定的倾向性:嘌呤核苷酸(A 与 G)的频率高达 60%~70%,而 C 的出现频率小于 17%。与原核生物 mRNA 相比,后者该位置上的核苷酸多为 U。据推测,可能是此位置上的核苷酸与终止作用的调节有关。

对许多编码细胞因子(如生长因子)的 mRNA 以及为癌基因编码的 mRNA 3'-UTR 序列的分析,发现其中包含着富含 UA 的保守序列,是由几个相间分布的 UUAUUUAU 八核苷酸序列组成。若除去这段序列,mRNA 的稳定性明显提高。可见 UA 序列是抑制翻译作用的元件,其调控特点是随着它在 3'-UTR 中拷贝数的增加,对翻译的抑制效率也提高。估计 UA 抑制翻译的机制可能是阻遏了起始阶段核糖体复合物形成前的某一过程。

3' poly(A) 尾的功能不仅是对 mRNA 由核内向细胞质运转时具有保护作用,而且对 mRNA 的稳定性和翻译效率还具有调控作用。有关植物、动物如爪蟾卵母细胞的体内实验,还有网织红细胞抽提物这样的高活性体外翻译体系中,都已观察到 mRNA poly(A) 尾结构与翻译效率之间的直接关系。

许多动物在卵母细胞中含有大量编码不同蛋白质的 mRNA,通常在卵子成熟前或受精前是不翻译的。这些 mRNA 原来多只有一个长 20~40 个腺苷酸尾,在卵细胞成熟或受精后胚胎发育过程中的适当时期,通过对一个外部信号的反应,才会将长上百个的 A 残基重新加到这些 mRNA 短的 poly(A) 尾上,然后翻译(图 15-18)。这种起调节作用的加尾是在细胞质中进行的,与核内 mRNA 前体的

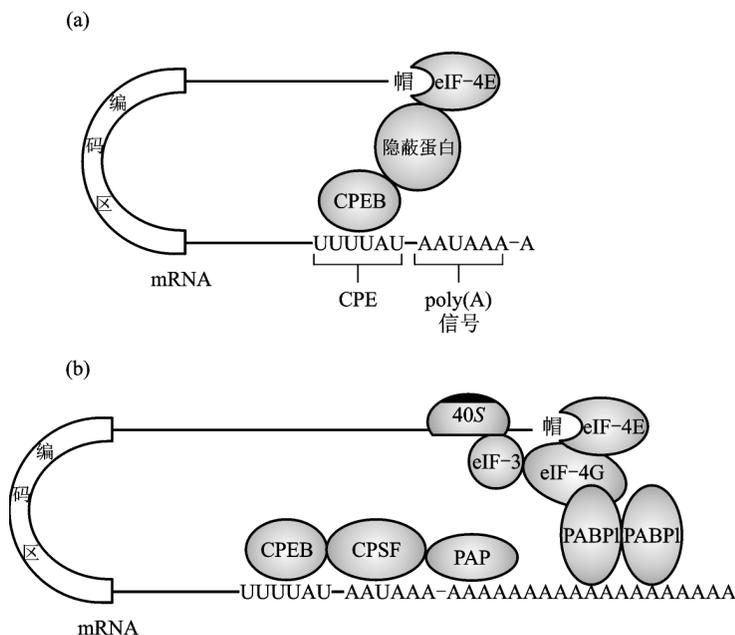


图 15-18 细胞质聚腺苷酸作用与翻译起始调控模式(引自 Lodish 等, 2004)

(a) 翻译暂停 (b) 翻译进行

加尾过程不同。爪蟾卵子细胞质中 mRNA 加尾的近期研究显示,细胞质中 mRNA 添加 poly(A)尾必须要有 AAUAAA(注:此序列是核内 RNA 前体加尾时也需要信号)以及细胞质聚腺苷酸作用元件(cytoplasmic polyadenylation element, CPE)。CPE 这一元件富含 U, 可被高度保守的 CPE-结合蛋白(CPE binding protein, CPEB)结合,此蛋白含有一个 RNA 识别结构域(RNA recognition motif RRM)和一个锌指功能域。没有刺激信号时,CPE 上的 CPEB 还与一种隐蔽蛋白质(maskin)相互作用,而后者又结合在与 5'-帽子相关的翻译起始因子 eIF-4E 上[图 15-18 a]。于是,使得 eIF-4E 无法与其他翻译起始因子以及核糖体 40S 小亚基发生相互作用,翻译起始受阻。当爪蟾卵母细胞成熟时,由于孕酮激素的影响,激活了可使 CPEB 磷酸化的激酶,隐蔽蛋白从被磷酸化的 CPEB 上释放;随即卵裂聚腺苷酸专一因子(cleavage and polyadenylation specificity factor, CPSF)以及 poly(A)聚合酶与 mRNA 结合。poly(A)聚合酶催化下 mRNA 的 poly(A)尾加长,胞质 poly(A)结合蛋白[cytoplasmic poly(A)-binding protein, PABP]就会结合到此延长的 poly(A)尾上,于是,mRNA 翻译起始所必需的全部参与物都能稳定地相互作用,启动翻译[图 15-18 b]。

### 15.3.3 翻译起始因子的可逆磷酸化与翻译调控

翻译因子的磷酸化修饰直接关系到翻译作用的激活和抑制。eIF-4E 是识别和结合 mRNA 5'-m<sup>7</sup>G 帽子结构的翻译起始因子,在哺乳动物中,此因子是由  $\alpha$ 、 $\beta$ 、 $\gamma$  3 个亚基组成,其中相对分子质量最小的  $\alpha$ -eIF-4E ( $2.4 \times 10^4$ ) 能直接结合到 mRNA 5'-m<sup>7</sup>G 帽子上,所以又称帽子结合蛋白; $\beta$  亚基相对分子质量  $4.6 \times 10^4$ ,是依赖于 RNA 的 ATP 酶,为 mRNA 与 40S 亚基结合时所必需;至于相对分子质量  $2.20 \times 10^5$  最大亚基的确切作用则有待进一步研究,推测在 eIF-3 和 40S 核糖体亚基相互作用中此 p220 亚基可能为 RNA 与主要蛋白的彼此结合提供静电接触。

eIF-4F 在 mRNA 翻译中的重要调控作用是通过其亚基的可逆磷酸化来实现的。对 eIF-4F 的 3 个亚基先采用 m<sup>7</sup>G 柱亲和层析的富集,再进行双向电泳和蛋白质染色,以测定它们的磷酸化状态。实验结果显示:静止期细胞当被胰岛素激活后,蛋白质的生物合成速度加快,而此时 eIF-4F 的  $\alpha$  和  $\gamma$  亚基的磷酸化作用增加;但当对细胞进行热激处理,或者是细胞处于生长周期的有丝分裂相时,蛋白质合成受到抑制,同时 eIF-4F 的  $\alpha$  亚基出现去磷酸化作用。

关于 eIF-4E 的磷酸化作用,从一些实验情况看,可能是有助于刺激 eIF-4F 的 3 个亚基形成复合物;或是促进 eIF-4B、eIF-4A 与 eIF-3 组装成更高级的复合物,因为这类复合物对翻译起始因子识别 mRNA 更为重要,可加快翻译的起始效率。

与 eIF-4F 磷酸化的激活效应不同,eIF-2 的磷酸化对翻译起始具有抑制作用。eIF-2 也是由 3 个相对分子质量不同的亚基组成。在有 GTP 的情况下,介导核糖体 40S 小亚基与甲硫氨基化的起始 tRNA (Met-tRNA) 特异结合,起始翻译(详见第 3 章)。然而,eIF-2 的磷酸化会抑制它的再利用。因为磷酸化的 eIF-2 对 GDP 和 eIF-2B 有很高的亲和力,从而便阻抑了 eIF-2 参与前述复合物的形成而再循环。

## 15.4 真核生物基因翻译后调节

### 15.4.1 新生肽链的剪接

翻译不是基因表达过程的最后一步。从 mRNA 翻译新合成的肽链多无活性,还必须在细胞质中加工和修饰才具有功能。新生肽链通过蛋白酶催化进行剪接加工,除去非功能片段。这些切割反应或是从 N 末端和(或)C 末端切除一小段,并折叠成有活性的蛋白分子,或是把含有多蛋白的肽链切割成数个有活性的成熟片段。

### (1) 多肽链末端的切割和多蛋白的水解加工

多肽链末端的裂解加工常见于分泌型多肽,如胰岛素多肽链的翻译后剪接加工。胰岛素是由胰岛的 B 细胞合成,最初在内质网膜上合成的肽链由信号肽、A、B 和 C 四个片段组成,这一肽链称为前胰岛素原(preproinsulin)。信号肽段在内质网中被信号肽酶切除,产生了胰岛素原(proinsulin)。胰岛素原进入高尔基体中,被 PC3 和 PC2 两种转变酶(convertase)从 C 片段两端切割,除去 C 片段。剩下的 A 和 B 片段通过二硫键结合,成为具有活性的胰岛素(insulin)(图 15-19)。

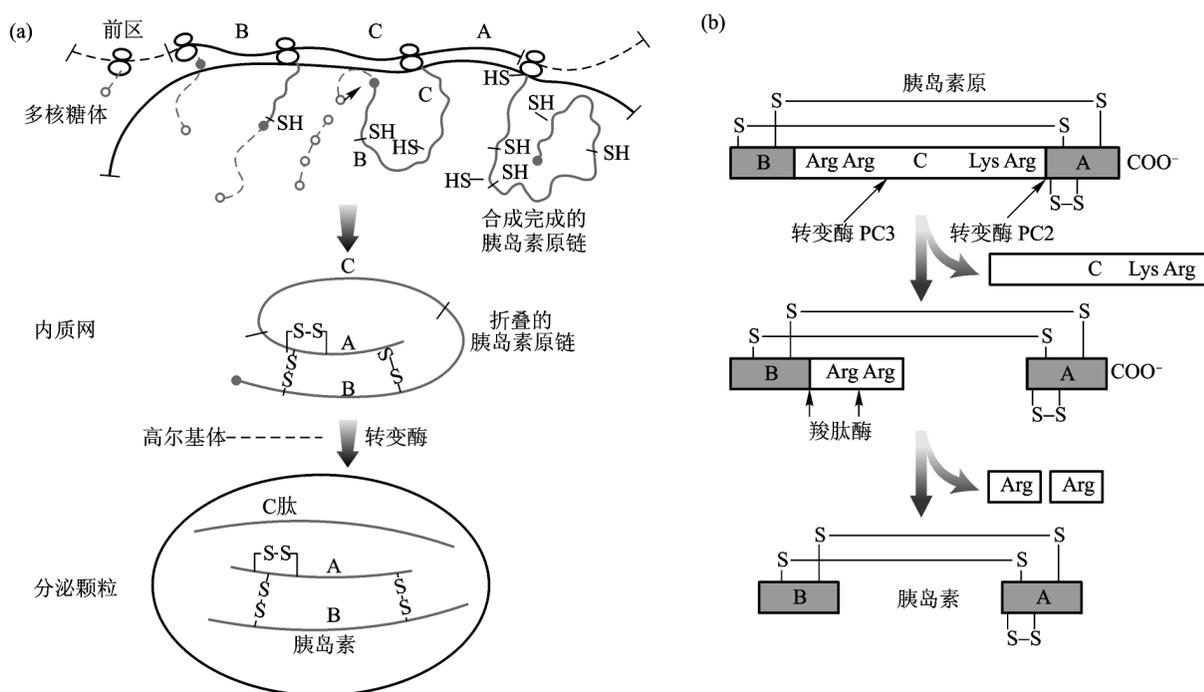


图 15-19 前胰岛素原加工成胰岛素示意图

有些多肽链在形成之初就是多蛋白肽链,经过剪接加工,才分解成几个各具不同功能的蛋白质。有几种病毒就是以这种方式利用小的基因组,编码多种蛋白质,几个基因共用一个启动子和一个终止子。真核生物中,多蛋白肽链并不罕见。脊椎动物中,有些肽类激素就有这种情况,如黑色素原(proopiomelanocortin, POMC)是在垂体前叶中合成的多蛋白肽链,其中至少包含 10 种不同的肽激素,经蛋白酶切割而被释放出来。只是这些肽激素的肽链序列彼此有所重叠,故酶切反应需分别进行,它们不能同时产生。在不同细胞中切割方式也有所不同。

### (2) 内含肽剪接

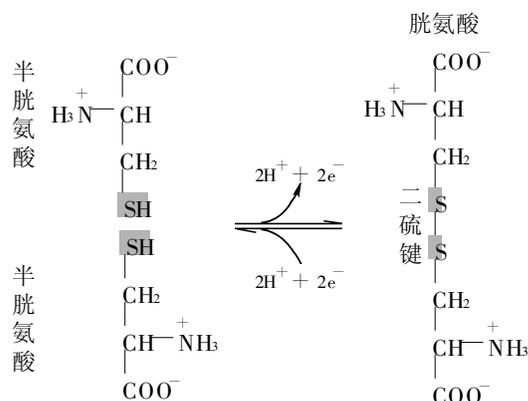
内含肽剪接(intein splicing)是指类似于前体 mRNA 剪接的修饰。内含肽(intein)是蛋白质肽链的内部片段,长 300~600 个氨基酸不等,在肽链剪接中被切除。剩下的肽片段称为外显肽(extein),拼接成成熟蛋白质。由于这个剪接过程与前体 RNA 的剪接极为相似,所以被称为蛋白质剪接(protein splicing)。1990 年首先在啤酒酵母中发现了内含肽。目前已知肽链中具有内含肽的蛋白质有百余种之多,遍布于真细菌、古细菌和真核生物中,而且在真核生物的核和细胞器蛋白质肽链中均有所发现。

内含肽的剪接不需要消耗能量,是通过一系列键的重排实现的,为一种自我催化反应。在所有已知的例子中,内含肽的剪接位点都是相似的,第一个氨基酸都是半胱氨酸,也有的是丝氨酸;最后两个氨基酸大都是组氨酸后接天冬氨酸。下游外显肽的第一个氨基酸常是半胱氨酸或丝氨酸或苏氨酸。内含肽序列中的其他氨基酸也是保守的,目前多认为这些保守氨基酸参与了剪接过程。

内含肽还有一个突出的特性,即有些内含肽具有归巢内切核酸酶(homing endonuclease)活性,归巢内切核酸酶切割靶 DNA,产生一个为该内含肽编码的 DNA 序列可插入的位点,从而将一个原来无内含肽编码序列的基因转变为含有内含肽编码的基因。蛋白质剪接与归巢内切核酸酶活性是彼此独立的。

### 15.4.2 新生肽链的化学修饰

一般,蛋白质的氨基酸有 20 种,但在成熟蛋白质中常存在有非常规氨基酸,这些氨基酸是在翻译后经化学修饰形成的,从而增加了氨基酸的种类,以适应蛋白质功能的需要。例如胱氨酸,在基因中无此密码子,但在肽链中可由二硫键将两个半胱氨酸连接成胱氨酸:



最简单的化学修饰类型是氨基酸的侧链或多肽链末端氨基酸的氨基或羧基加接小的化学基团,如乙酰基、甲基或磷酸基。修饰的方式有多种,如胶原中的脯氨酸和赖氨酸的羟基化。目前已发现有 150 余种修饰氨基酸,每一种都有特定的修饰方式。这些修饰往往对蛋白质的精确活性起决定作用。如组蛋白乙酰化、甲基化对确定染色体的精细结构具有重要作用。有些蛋白质的磷酸化与信号转导直接相关。

在蛋白质的修饰中最为复杂的是糖基化(glycosylation)。糖基化是指多肽链上连接多糖链,常见的有两种类型:O-连接的糖基化和 N-连接的糖基化。前者是把糖基链连接到肽链中丝氨酸或苏氨酸的羟基上;后者是糖链连接到天冬酰胺的氨基上。蛋白质的糖基化不仅关系到信号分拣,而且在蛋白质正确折叠、增加蛋白质的稳定性、抵御酶降解以及参与细胞识别等方面均具有重要作用。

### 15.4.3 肽链的折叠

有功能的成熟蛋白质分子具有一定的三维构型和构象。新生肽链经过上述几种方式修饰后,还必须进行正确的折叠才能成为有功能的蛋白质,折叠错误则无功能,甚至可引起疾病。无论是原核细胞还是真核细胞,都含有一类能使蛋白质肽链正确折叠的蛋白质,这类蛋白质称为分子伴侣(chaperone)。在蛋白质折叠和组装过程中,分子伴侣能够防止多肽链的链内和链间错误折叠或聚集作用,且还能破坏多肽链中已形成的错误结构,但自身不参加最终产物的组成。在真核生物中,蛋白质折叠主要依赖于 Hsp70 及其同系物,因为它们可以在 mRNA 翻译的同时进行肽链折叠;另外,还有一种 Hsp40 辅分子伴侣(co-chaperone)可以加速这一折叠过程。

### 15.4.4 蛋白质更换

在真核生物中,一个细胞内各种蛋白质的数量,既取决于新生肽链的合成速率,又取决于它们存活的寿命,所以在一定时期内,细胞内的一些蛋白质降解为氨基酸,用以调节其特定蛋白质的数量。细胞内的蛋白质除可以通过溶酶体水解外,还具有蛋白酶解(proteolysis)的专门途径。这一专门的蛋

白酶解活动与细胞的生理活动、细胞周期以及细胞分化等过程密切相关。在真核细胞内,注定要降解的蛋白质首先经过专一性的蛋白质泛素化酶(ubiquitin enzyme)催化,和泛素(ubiquitin)分子共价结合,挂上“清除标签”。短寿命的蛋白质常具有一段短的氨基酸序列,它是蛋白质降解的信号序列。至于那些变性的或折叠错误的蛋白质,以及含有氧化氨基酸或其他异常氨基酸的蛋白质,都可被依赖泛素的蛋白质酶的酶解系统所识别,并降解。

在真核细胞的细胞质溶质中,大多数蛋白质降解是由一种大的蛋白酶复合物来完成。蛋白质的这种降解场所是由一些保守性蛋白质所组成的特殊装置,称为蛋白酶体(proteasome),细胞内含有许多蛋白酶体。

蛋白酶体是一个大的多亚基结构,沉降系数为 36S。外形类似有底有盖的筒状体。中空的筒状体为 20S,由一些不同的蛋白酶所组成,是蛋白酶体的核心。进入筒状体内腔的蛋白质被蛋白酶降解成短肽,然后释放回细胞质溶质,再进一步降解成氨基酸单体,在蛋白质合成中又可被重新利用。

## 15.5 真核生物基因表达中的 RNA 调节

### 15.5.1 RNA 干扰

近年来发现,一些 RNA 在基因表达调节中扮演着重要角色。它们以阻止 mRNA 翻译、降解 mRNA,或是通过控制 mRNA 表达的启动子以及发生转录后沉默等方式来抑制其同源基因(homologous gene)的表达。这种 RNA 有效阻断同源基因表达的现象即称为 RNA 干扰(RNA interference, RNAi)。

RNA 干扰现象首先是在研究线虫(*C. elegans*)时发现的,即用正义链也能有效地抑制同源基因的表达。1998年,Fire和 Mello发现,双链 RNA 抑制同源基因表达的能力比反义链或正义链更为有效,而且每个细胞只有几个双链 RNA(dsRNA)分子就能产生抑制效应。向线虫生殖腺中注入双链 RNA,可导致整个虫体中的同源基因发生沉默,而且可传到下一代。关于双链 RNA 抑制基因表达的作用在果蝇中也存在,如把双链 RNA 注入到果蝇胚胎中,能够强烈地抑制同源基因的表达,而单链 RNA 的作用则较微弱。后来,类似的现象在动物、真菌、植物细胞中陆续发现。由此看来, RNA 干扰现象在生物界中普遍存在。推测,这可能是一种古老的保护机制,既可用于发育调节,又可作为防御某些病毒入侵的手段。

由于双链 RNA 对同源基因破坏的专一性,几乎能使体内任何特定基因发生表达沉默。因而,与从基因组内打断编码序列的方法相比,就更为简单。作为一种实验技术, RNA 干扰已成为研究基因功能的有力手段之一。故 2006年诺贝尔评审委员会指出, RNA 干扰机制将有望应用于临床医学和农业众多领域,用来开发针对病毒感染、心血管疾病和癌症等的新疗法。近来有动物研究结果显示,可利用 RNA 干扰机制使高血脂基因“沉默”。

### 15.5.2 小 RNA 在基因表达中的调节作用

#### (1) 微 RNA 在基因表达中的调节作用

与一些调节蛋白(阻遏物和激活因子)一样,一些 RNA 在细胞中也具有调节基因表达的作用。它们的作用方式是:通过碱基配对,与目的核苷酸序列互补形成双链区,直接阻止后者功能的发挥,或是与目的核苷酸序列中的某一部分形成双链区,以使后者的构象发生改变,抑制其发挥作用。这类由 RNA 介导的调节作用,其目的核苷酸序列是同源目的 mRNA 的一部分。二者相互作用,发生了二级结构变化,形成一个双链 RNA 的发夹结构,阻遏目的序列发挥作用。

在动物和植物细胞中,有许多很小的 RNA 分子,它们由 22个左右的核苷酸组成,称为微 RNA

(micro RNA, miRNA)。微 RNA 是通过与目的 mRNA 序列中的部分碱基序列配对,调节基因的表达。在秀丽隐杆线虫(*Caenorhabditis elegans*)中发现了 micro RNA——lin 4 RNA,它与 lin 14 mRNA 相互作用,使后者的表达受阻。lin 14 基因调节幼虫的发育,但它的表达受 lin 4 的控制。lin 4 RNA 中含有可与 lin 14 mRNA 3'-UTR 某些区段中的同源序列互补,从而阻遏 lin 14 转录后的翻译(图 15-20)。

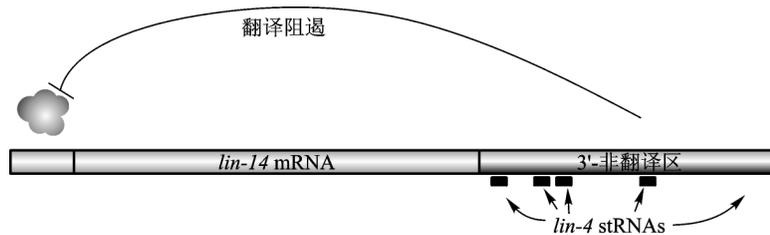


图 15-20 lin 4 mRNA 结合 lin 14 mRNA 3'-UTR 上的多个部位(仿自 M. T. Mamus 等, 2003)  
sRNA, small temporal RNA(小时序 RNA)

lin 4 RNA 是从带有发夹结构的 lin 4 前体 RNA 加工而来。在秀丽隐杆线虫基因组中约有 55 个基因为这类 mRNA 编码。发育时,它们具有各种不同的表达类型,但都是作为基因表达的调节子。

在哺乳动物中已发现许多秀丽隐杆线虫 mRNA 同源物;在植物中也有类似情况,如拟南芥(*Arabidopsis*)的 16 个 mRNA 中有 8 个完整地存在于水稻中。由此可见,这种调节机制广泛存在于真核生物中。

mRNA 产生的机制十分恒定,是由非蛋白编码基因转录而来的大前体物加工而成。这种转录物有 70~90 个核苷酸长,含有形成发夹结构的序列,因而可形成双链区。而这个双链区恰好成为双链 DNA 切丁酶(Dicer)的靶子,Dicer 酶将转录本加工,产生了具有活性的 mRNA。

### (2) 小干扰 RNA 对基因表达的抑制

RNA 干扰发现不久,人们进一步在线虫、拟南芥、链孢霉、衣藻,以及果蝇等真核生物中鉴定出与基因沉默有关的基因。这些基因表达的抑制都是发生在细胞质内,因而特称为转录后基因沉默(post transcriptional gene silence, PTGS)。经序列分析和分子杂交鉴定表明,这类沉默现象和 RNA 干扰一样都是由一种很小的双链 RNA 造成的。这种小双链 RNA 只有 21~23 个核苷酸长,3'端还常有 2 个核苷酸单独伸出。由于它们能够与同源 mRNA 互补配对,进而诱导相关的酶降解其所互补的 mRNA。因此,就把这些极小的 RNA 分子称为小干扰 RNA(small interference RNA, sRNA)。

美国冷泉港实验室的科学家用 dsRNA 来转染果蝇细胞,然后从细胞提取物分离出具有核酸酶活性的部分,他们发现,这种核酸酶在细胞外能特异地降解与 dsRNA 有同源序列的 mRNA。如果在细胞外反应体系加入 mRNA 之前,先用微球菌核酸酶(micrococcal nuclease)(可降解 DNA 或 RNA)处理所分离得到的核酸酶(含有 sRNA),就不会出现 RNA 干扰现象。用 DNA 酶处理则不影响 RNA 干扰。这些结果说明 sRNA 在目标识别中的作用。

### 15.5.3 RNA 干扰的机制

近几年,有关 RNA 干扰的研究迅速展开。尽管其中还有许多细节尚不清楚,但根据对一些不同物种中 RNA 干扰的生物化学和遗传学研究,已对这一反应过程的机制提出了一个基本模式(图 15-21)。鉴定出的几种与 RNA 干扰相关的蛋白质和蛋白质复合物,以及反应中介物,为模型建立提供了实验依据。如 Dicer 酶,这是一种类似核苷酸酶 III 的生物酶,它负责切割双链 RNA,产生大约 23 个核苷酸残基的双链片段。还有一种可和 sRNA 组装成包含有多种蛋白质的 RNA 沉默复合体(RNA-induced silencing complex, RISC),也是 RNA 干扰过程中必需的一种成分。

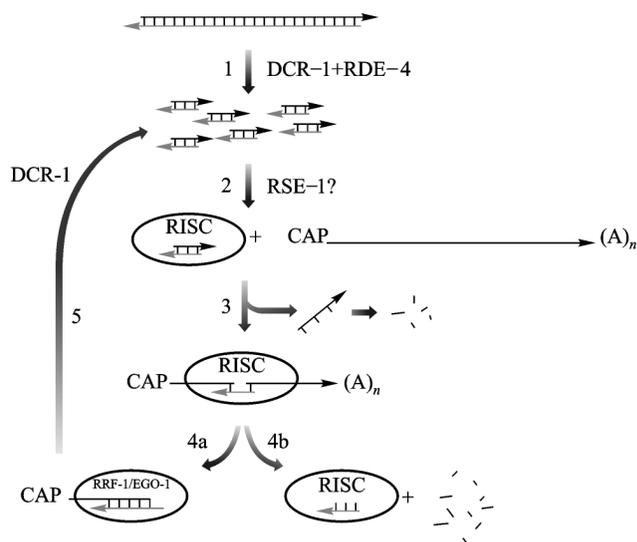


图 15-21 RNA 的作用机制模型(引自 Ketting 等, 2003)

RNA 干扰的第一步是由 *Dicer* 将双链 RNA 切割成约 23bp 的片段。1999 年首先在植物转录后沉默现象中发现 *Dicer* 对 dsRNA 的加工。此后, 又陆续在许多种生物中发现 dsRNA。在果蝇、秀丽隐杆线虫的体内、体外系统中, 更进一步确证了 sRNA 的产生是 RNA 干扰最基本的一个环节。此外, 根据线虫体内系统的研究, 在 *Dicer* 切割 dsRNA 的过程中, 还需要另一种蛋白质——受体破坏酶 (receptor destroying enzyme RDE-4) 的参与, 以促使对 dsRNA 的有效加工。加工产生的 sRNA 随后掺入 RISC 复合体。二者一旦进行组装, 复合体中的核酸酶成分就会在 ATP 作用下催化 dsRNA 的变性和置换反应。单链 RNA 的出现, 激活 RISC 复合体, 激活后的复合体被引导到与其中 sRNA 互补的 mRNA 序列。因 RNA 的置换反应, 原 sRNA 的反义链便可与 mRNA 中相应的序列互补, 或降解该 mRNA, 或抑制其翻译 (图 15-21) (图中 DCR = *Dicer*)。在此模型中, 后期处理路线的选择基本上取决于 sRNA 与同源 mRNA 之间的匹配程度。如果二者完全互补, 则发生降解; 如果匹配程度不很合适, 则在很大程度上是抑制翻译。有时候, 在同源 mRNA 被切割之后, 还可发生两种情况: 一是通过 RNA 依赖的 RNA 聚合酶的作用, 使 sRNA 延长, 导致 dsRNA 再次进入 RNA 途径。二是受到切割的 mRNA 进一步降解。而完成此任务后的 RISC 复合物则参与另一轮 mRNA 的降解循环。

从上述有关 RNA 机制的描述中, 表明 RNA 反应的几个基本特点是: ① 在 RNA 过程中, 目的基因的内源序列没有改变, 也就是说 RNA 并不会造成稳定的遗传变化, 但有些 RNA 的效应却可以传递一两个世代。② 目的 mRNA 的衰减是发生在细胞质中, 并不影响核内的前体 mRNA。在实际观察中也表明, RNA 与位于染色体上的基因并没有交叉反应。这就有力地说明 RNA 是通过增加 mRNA 特异性的速率周转, 来行使对基因转录后的调控作用。这种现象在植物中表现得尤为突出, 使人们认识到 RNA 与转录后的基因沉默间的联系。③ RNA 沉默的一个突出特征是作用的效率极高。少量的 dsRNA 就足以影响一个大的 mRNA 库, 促使目标基因彻底关闭。④ RNA 效应还可以在生物个体内扩散, 如秀丽隐杆线虫中, 把 dsRNA 注入生殖腺, 其效应可扩散到虫体全身。

## 思考题

1. 请列表说明真核与原核基因表达的差异。
2. 为什么说核小体修饰与 DNA 甲基化是表观遗传的基础?
3. 基因转录水平的调节是真核基因表达中最为关键的步骤, 你同意吗? 为什么?

4. 选择性剪接、**RNA**编辑等转录后的特殊调控有什么生物学意义?
5. 如果一个基因为父本印记,一个受到影响的男孩,是从哪个亲本遗传了这一突变基因?
6. 在真核基因表达中辅激活物具有什么作用?
7. 请简述染色质重塑在真核基因表达中的作用。
8. **mRNA**中的非编码序列与 **mRNA**翻译调控有何关系?
9. 在真核基因调控中一些小 **RNA**分子发挥了什么作用? 其机制如何?